On finite sequences satisfying linear recursions

Noam D. Elkies[1]

May, 2001

corrected April, 2002 per referee's report

Abstract. For any field k and any integers m, n with $0 \leqslant 2m \leqslant n+1$, let W_n be the k-vector space of sequences $(x_0, \ldots, x_n)$, and let $H_m \subseteq W_n$ be the subset of sequences satisfying a degree-m linear recursion, i.e. for which there exist $a_0, \ldots, a_m \in k$ such that subset of sequences satisfying a degree-m linear recursion — that is, for which there exist $a_0, \ldots, a_m \in k$, not all zero, such that

$$\sum_{i=0}^{m} a_i x_{i+j} = 0$$

holds for each $j = 0, 1, \ldots, n-m$. Equivalently, H_m is the set of $(x_0, \ldots, x_n)$ such that the $(m+1) \times (n-m+1)$ matrix with (i, j) entry x_{i+j} $(0 \leqslant i \leqslant m,\ 0 \leqslant j \leqslant n-m)$ has rank at most m. We use elementary linear and polynomial algebra to study these sets H_m. In particular, when k is a finite field of q elements, we write the characteristic function of H_m as a linear combination of characteristic functions of linear subspaces of dimensions m and $m+1$ in W_n. We deduce a formula for the discrete Fourier transform of this characteristic function, and obtain some consequences. For instance, if the $2m+1$ entries of a square Hankel matrix of order $m+1$ are chosen independently from a fixed but not necessarily uniform distribution μ on k, then as $m \to \infty$ the matrix is singular with probability approaching $1/q$ provided $\|\widehat{\mu}\|_1 < q^{1/2}$. This bound $q^{1/2}$ is best possible if q is a square.

Introduction

Fix a field k. For any integers m, n with $0 \leqslant 2m \leqslant n+1$, let W_n be the k-vector space of sequences $(x_0, \ldots, x_n)$, and let $H_m \subseteq W_n$ be the subset of sequences satisfying a degree-m linear recursion, that is, for which there exist $a_0, \ldots, a_m \in k$, not all zero, such that

$$\sum_{i=0}^{m} a_i x_{i+j} = 0 \tag{1}$$

holds for each $j = 0, 1, \ldots, n-m$. Equivalently, H_m is the set of $(x_0, \ldots, x_n)$

[1] Supported in part by the Packard Foundation. e-mail address: elkies@math.harvard.edu

such that the $(m+1)\times(n-m+1)$ Hankel matrix[2]

$$\begin{pmatrix} x_0 & x_1 & \cdots & x_{n-m} \\ x_1 & x_2 & \cdots & x_{n-m+1} \\ \vdots & \vdots & & \vdots \\ x_m & x_{m+1} & \cdots & x_n \end{pmatrix} \tag{2}$$

has rank at most m.[3] Now linear recursions on *infinite* sequences $\{x_i\}_{i\in\mathbf{Z}}$ are known to correspond to polynomials in the shift operators $T^{\pm 1} : \{x_i\} \mapsto \{x_{i\pm 1}\}$, modulo multiplication by powers of T. This approach does not work so nicely for finite sequences, because T and T^{-1} push x_0 and x_n off the edge. We propose to remedy this problem at $T = 0, \infty$ by homogenizing: instead of polynomials in $T^{\pm 1}$, use homogeneous polynomials in two variables Y and Z that act on W_n as the right and left truncation maps to W_{n-1}. We shall see that this approach yields a clean account of linear recursions and the subsets H_m in the space W_n, which itself will be identified with the dual of the space V_n of homogeneous polynomials of degree n in Y and Z.[4]

In the present paper we develop this account using elementary linear and polynomial algebra. When k is a finite field of q elements, we also write the characteristic function of H_m as a linear combination of characteristic functions of linear subspaces of dimensions m and $m+1$ in W_n. We deduce a formula for the discrete Fourier transform of this characteristic function, and obtain some consequences. For instance we obtain a new proof that $\#H_m = q^{2m}$. We further show that if the $2m+1$ entries $x_0, \ldots, x_{2m}$ of a square Hankel matrix

$$\begin{pmatrix} x_0 & x_1 & \cdots & x_m \\ x_1 & x_2 & \cdots & x_{m+1} \\ \vdots & \vdots & & \vdots \\ x_m & x_{m+1} & \cdots & x_{2m} \end{pmatrix} \tag{3}$$

of order $m+1$ are chosen independently from a fixed but not necessarily uniform distribution μ on k, then as $m \to \infty$ the the matrix is singular with probability approaching $1/q$ provided the Fourier transform of μ has l_1 norm less than $q^{1/2}$. This bound is best possible if q is a square: if μ is the uniform distribution

[2] For more background on Hankel matrices (matrices with entries constant on NE-SW diagonals), and the closely related Toeplitz or "persymmetric" matrices (with entries constant on NW-SE diagonals), see for instance [4]. These matrices arise in diverse mathematical contexts; see for instance [1, 3] and the references in [4]. In our setting, Hankel matrices are more natural than Toeplitz ones, but our results on rank distribution, culminating in Thm. 2, apply equally well to matrices of either Hankel or Toeplitz type.

[3] I thank Joe Harris for the geometric observation that H_m consists of the lines through the origin coming from the points $(x_0 : x_1 : \cdots : x_n)$ lying on the m-th secant variety of the rational normal curve $(\xi^n : \xi^{n-1}\eta : \cdots : \eta^n)$ in n-dimensional projective space over k. We shall not need this formulation here, but it arises naturally in an arithmetic application of H_m [3].

[4] As an added benefit, the whole structure inherits a $\mathrm{GL}_2(k)$ structure from the action of $\mathrm{GL}_2(k)$ by linear substitutions on Y, Z. But this, too, is not needed for the present paper.

on ck_0, where k_0 is a quadratic subfield of k and $c \in k^*$ is arbitrary, then $\|\widehat{\mu}\|_1 = q^{1/2}$ but the probability is $q^{-1/2}$. It seems reasonable to conjecture that for any μ the matrix (3) is singular with probability $\to 1/q$ as long as μ is supported on a set of at least two elements not contained in ck_0 for any proper subfield k_0 of k.

The spaces V_n, W_n and some linear algebra

Basic notions and lemmas. Fix a field k. For each integer $n \geq -1$, let V_n be the vector space of dimension $n+1$ over k consisting of bivariate homogeneous polynomials

$$P(Y,Z) = \sum_{i=0}^{n} a_i Y^i Z^{n-i} \tag{4}$$

of degree n. Let W_n as the dual of V_n. We identify W_n with the space of sequences $(x_0, \ldots, x_n)$ by regarding such a sequence as the linear functional

$$\sum_{i=0}^{n} a_i Y^i Z^{n-i} \mapsto \sum_{i=0}^{n} a_i x_i \tag{5}$$

on V_n. Note that we allow V_{-1} and W_{-1}, each of which is the zero space, but not V_n, W_n for $n < -1$.

If $m \geqslant 0$ and $n+1 \geqslant m$, polynomial multiplication $V_m \times V_{n-m} \to V_n$ gives for each $Q \in V_m$ a linear map $M_n(Q) : V_{n-m} \to V_n$ defined by

$$M_n(Q) : P \mapsto PQ \quad (P \in V_{n-m}). \tag{6}$$

Our reason for identifying the space W_n of sequences $(x_0, \ldots, x_n)$ with the dual of V_n is the following observation:

Lemma 1. *Suppose $Q \in V_m$ is the polynomial $\sum_{i=0}^{m} a_i Y^i Z^{m-i}$. Then the adjoint of $M_n(Q)$ is the linear map $M_n^*(Q) : W_n \to W_{n-m}$ taking any $(x_0, \ldots, x_n)$ to the sequence of length $n-m+1$ whose j-th term is*

$$\sum_{i=0}^{m} a_i x_{i+j} \tag{7}$$

for each j with $0 \leqslant j \leqslant n-m$.

Proof: We show the equivalent dual statement: the linear map $M_n(Q)$ takes any polynomial

$$P(Y,Z) \sum_{j=0}^{n-m} b_j Y^j Z^{n-m-j}$$

in V_{n-m} to the polynomial $PQ \in V_n$ whose $Y^r Z^{n-r}$ coefficient is $\sum_{i+j=r} a_i b_j$ for each r with $0 \leqslant r \leqslant n$. But this is immediate from the expansion of PQ. ■

Thus H_m is the union of $\ker M_n^*(Q)$ over all nonzero $Q \in V_m$.

Of course that union is not disjoint, but as long as $2m \leqslant n+1$ we shall describe the intersection of $\ker M_n^*(Q_1)$ and $\ker M_n^*(Q_2)$ for any Q_1, Q_2 of degree at most m, see Lemma 4 below. We first establish some further basic properties:

Lemma 2. *i) For any $Q, Q' \in V_m$ and n such that $n+1 \geqslant m \geqslant 0$ we have*

$$M_n^*(Q+Q') = M_n^*(Q) + M_n^*(Q'). \tag{8}$$

ii) For any $Q_1 \in V_{m_1}$, $Q_2 \in V_{m_2}$, and n such that $m_1, m_2 > 0$ and $n+1 \leqslant m_1 + m_2$, we have

$$M_n^*(Q_1 Q_2) = M_{n-m_2}^*(Q_1) \circ M_n^*(Q_2) = M_{n-m_1}^*(Q_2) \circ M_n^*(Q_1). \tag{9}$$

iii) For any nonzero $Q \in V_m$ and any $n \geq m-1$, the map $M_n^*(Q)$ is surjective and its kernel has dimension m.

Proof: (i) This is the dual of the identity $M_n(Q+Q') = M_n(Q) + M_n(Q')$, which is just the distributive law $P(Q+Q') = PQ + PQ'$ for multiplication of homogeneous polynomials. (Alternatively, apply Lemma 1.)

(ii) Likewise this is the dual of the fact that multiplying a polynomial of degree $n - m_1 - m_2$ by $Q_1 Q_2$ is the same as multiplying it first by Q_2 and then by Q_1 or vice versa.

(iii) Since $k[Y, Z]$ has no zero divisors, $M_n(Q)$ is injective; thus $M_n^*(Q)$ is surjective, and its kernel has dimension $\dim W_n - \dim W_{n-m} = m$. ■

The ideal I_x. For $x = (x_0, \ldots, x_n) \in W_n$, define $I_x \subseteq k[Y, Z]$ as follows: any $Q \in k[Y, Z]$ is uniquely $\sum_{m=0}^{M} Q_m$ with each $Q_m \in V_m$; the subset I_x consists of those $\sum_{m=0}^{M} Q_m$ for which $(M_n^*(Q_m))(x) = 0$ for each $m \leqslant n+1$.

Lemma 3. *I_x is a homogeneous ideal in $k[Y, Z]$ for all $x \in W_n$.*

Proof: By definition $\sum_{m=0}^{M} Q_m \in I_x$ if and only if each $Q_m \in I_x$. So it is enough to check that $I_x \cap V_m$ is closed under addition for each m, and that $PQ \in I_x$ if $Q \in I_x \cap V_m$ and $P \in V_{m'}$ for some $m, m' \geq 0$. Each of these is vacuously true if $m > n+1$ or $m + m' > n+1$ respectively, and follows from part (i) or (ii) of Lemma 2 otherwise. ■

The main result of this section is the following partial description of I_x, stating in effect that it is approximated by a principal ideal as well as dimension considerations allow:

Proposition 1. *Suppose for some $x \in W_n$ that I_x contains a nonzero polynomial of degree at most $(n+1)/2$. Let m_0 be the smallest degree of such a polynomial. Then $I_x \cap V_{m_0}$ is 1-dimensional, say*

$$I_x \cap V_{m_0} = kQ_0 \tag{10}$$

for some nonzero $Q_0 \in V_{m_0}$. For each $m \leqslant n+1-m_0$,

$$I_x \cap V_m = (M_n(Q_0))\,(V_{m-m_0}). \tag{11}$$

Remark: In particular, it follows that $I_x \cap V_m$ has dimension $m - m_0 + 1$ for $m_0 \leqslant m \leqslant n + 1 - m_0$. This cannot hold once $m > n + 1 - m_0$, except in the trivial case $x = 0$, when $m_0 = 0$ and I_x is all of $k[Y, Z]$. Indeed suppose that $m_0 > 0$ and $m > n + 1 - m_0$. If $m > n + 1$ then $I_x \cap V_m = V_m$ has dimension $m + 1 > m - m_0 + 1$. If $m \leqslant n + 1$ then $I_x \cap V_m$ is the kernel of the linear map

$$V_m \to W_{n-m}, \quad Q \mapsto (M_n^*(Q))(x); \tag{12}$$

thus

$$\dim(I_x \cap V_m) \geqslant \dim V_m - \dim W_{n-m} = 2m - n, \tag{13}$$

which again exceeds $m - m_0 + 1$ since $m > n + 1 - m_0$. This is what we mean when we state that I_x approximates the principal ideal (Q_0) as well as dimension considerations allow.

To prove Prop. 1 we must first make good on our promise to describe intersections of the spaces $\ker M_n^*(Q)$. We do this in the next lemma, whose statement uses the greatest common divisor Q of two homogeneous polynomials Q_1, Q_2. This is defined only up to multiplication by k^*, but such scaling does not affect the space $\ker M_n^*(Q)$, so the choice of g.c.d. will not affect the result.

Lemma 4. *Let Q_1, Q_2 be nonzero polynomials in V_{m_1}, V_{m_2} respectively, with greatest common divisor Q. Then, for each $n \geqslant \max(m_1, m_2) - 1$,*

$$\ker M_n^*(Q_1) \cap \ker M_n^*(Q_2) \supseteq \ker M_n^*(Q), \tag{14}$$

with equality if and only if

$$n + 1 \geqslant m_1 + m_2 - \deg(Q). \tag{15}$$

Proof: If $x \in \ker M_n^*(Q)$ then x is in the kernel of both $M_n^*(Q_1)$ and $M_n^*(Q_2)$, because each of these linear maps factors through $M_n^*(Q)$ by part (ii) of Lemma 2. Thus x is in the intersection of the two kernels, whence (14) follows. It remains to establish the condition of equality.

Let $m = \deg Q$, and $m' = m_1 + m_2 - m$. By Lemma 2(iii), the codimensions in W_n of $\ker M_n^*(Q_1)$ and $\ker M_n^*(Q_2)$ are $n+1-m_1$ and $n+1-m_2$ respectively. Thus their intersection has codimension at most

$$(n + 1 - m_1) + (n + 1 - m_2) = (n + 1 - m) + (n + 1 - m'). \tag{16}$$

Hence if $m' > n + 1$ then this codimension is strictly less than the codimension of $\ker M_n^*(Q)$. Thus the condition $m' \leqslant n + 1$ is necessary for equality in (14).

We conclude the proof by showing that this condition is also sufficient. Let Q' be the least common multiple

$$Q' = Q_1 Q_2 / Q \tag{17}$$

of Q_1 and Q_2; this is a homogeneous polynomial of degree m'. Assuming that $m' \leqslant n + 1$, we may then consider $M_n^*(Q')$. We claim that

$$\ker M_n^*(Q_1) + \ker M_n^*(Q_2) = \ker M_n^*(Q'). \tag{18}$$

By duality, this claim is equivalent to

$$\mathrm{im}(M_n(Q_1)) \cap \mathrm{im}(M_n(Q_2)) = \mathrm{im}(M_n(Q')). \tag{19}$$

But this is just the statement that a polynomial in V_n is divisible by both Q_1 and Q_2 if and only if it is divisible by Q' — which is true because Q' is the least common multiple of Q_1 and Q_2. We thus have

$$\begin{aligned}
& \dim(\ker M_n^*(Q_1) \cap \ker M_n^*(Q_2)) \\
= \quad & \dim(\ker M_n^*(Q_1)) + \dim(\ker M_n^*(Q_2)) - \dim(\ker M_n^*(Q_1) + \ker M_n^*(Q_2)) \\
= \quad & \dim(\ker M_n^*(Q_1)) + \dim(\ker M_n^*(Q_2)) - \dim(\ker M_n^*(Q')). \qquad (20)
\end{aligned}$$

By Lemma 2(iii) again, this dimension equals

$$m_1 + m_2 - m' = m = \dim(\ker M_n^*(Q)). \tag{21}$$

Since we already know that $\ker M_n^*(Q_1) \cap \ker M_n^*(Q_2)$ contains $\ker M_n^*(Q)$, we conclude that these two spaces are equal. ■

Corollary. *Suppose $x \in W_n$. If I_x contains homogeneous polynomials Q_1, Q_2 whose least common multiple has degree at most $n+1$, then I_x contains* $\gcd(Q_1, Q_2)$. *In particular, this conclusion holds if* $\deg Q_1 + \deg Q_2 \leqslant n + 1$.

Proof: Under our hypotheses, x is contained in both $\ker M_n^*(Q_1)$ and $\ker M_n^*(Q_2)$, and the equality condition of Lemma 4 is satisfied. Therefore

$$x \in \ker M_n^*(Q_1) \cap \ker M_n^*(Q_2) = \ker M_n^*(\gcd(Q_1, Q_2)), \tag{22}$$

which is to say that I_x contains $\gcd(Q_1, Q_2)$ as claimed. ■

We can now easily prove Prop. 1. Suppose Q_1, Q_2 are nonzero polynomials in $I_x \cap V_{m_0}$. By the hypothesis of Prop. 1 we know $2m_0 \leqslant n+1$. The Corollary to Lemma 4 thus applies, and we find that I_x contains $\gcd(Q_1, Q_2)$. Unless Q_1, Q_2 are proportional, $\deg(\gcd(Q_1, Q_2)) < m_0$, which is impossible by the definition of m_0. Thus $I_x \cap V_{m_0}$ has dimension 1 as claimed. By the same Corollary, if $m \leqslant n + 1 - m_0$ and $Q \in I_x \cap V_{m_0} - \{0\}$ then $I_x \ni \gcd(Q_0, Q)$. Since again $\gcd(Q_0, Q)$ must have degree at least m_0, we conclude that Q is a multiple of Q_0. Since I_x is an ideal (Lemma 3), we already know that I_x contains all multiples of Q_0; thus $I_x \cap V_{m_0}$ consists of all degree-m multiples of Q_0, and we are done. ■■

It is thus natural to call Q_0 the *minimal linear recursion* satisfied by x. (Again Q_0 is defined only up to multiplication by k^*.) From Prop. 1 we deduce the following description of the degree m_0 of this minimal recursion:

Corollary. *If $x \in H_m$ for some $m \leqslant (n+1)/2$ then the degree of the minimal linear recursion satisfied by x equals the rank of the Hankel matrix (2) associated to x.*

Proof: Let m_0 be this minimal degree. The rank of (2) is $m + 1 - d$, where d is the dimension of the kernel of the action of this matrix on row vectors of length $m+1$. But Lemma 1 identifies this kernel with the space $I_x \cap V_m$ of degree-m recursions satisfied by x. Since $m_0 \leqslant m \leqslant (n+1)/2$, we may apply Prop. 1 to find that $d = m - m_0 + 1$. Thus m_0 is the rank of the Hankel matrix, as claimed. ■

The characteristic function of H_m

Decomposition into signed linear subspaces. We assume henceforth that k is a finite field of q elements. For integers m, n satisfying our customary condition $2m \leqslant n+1$, let $\mathbf{P}_m$ be the set of all subspaces of W_n of the form $\ker M_n^*(Q)$ for some nonzero $Q \in V_m$. (By Lemma 4 and part (iii) of Lemma 2, $\ker M_n^*(Q_1) = \ker M_n^*(Q_2)$ if and only if Q_1, Q_2 are proportional; thus $\mathbf{P}_m$ consists of

$$\frac{\#(V_m - \{0\})}{\#(k^*)} = \frac{q^{m+1}-1}{q-1} \tag{23}$$

subspaces. We note for later use that this formula remains valid if we allow $m = -1$, when $\mathbf{P}_m$ is empty.) Recall that we defined H_m as the set of $x \in W_n$ satisfying a recursion of degree m, and noted that H_m is thus the union of all the subspaces in $\mathbf{P}_m$. We further noted that this union is not disjoint, and thus that χ_{H_m}, the characteristic function of H_m, is not simply the sum of the characteristic functions of the subspaces in $\mathbf{P}_m$. However, by Lemma 4, the intersection of any two subspaces in $\mathbf{P}_m$ is again the kernel of $M_n^*(Q)$ for some nonzero homogeneous Q of degree $\leqslant m$, and more generally if $m_1, m_2 \leqslant m$ then the intersection of any subspace in $\mathbf{P}_{m_1}$ with any subspace in $\mathbf{P}_{m_2}$ is itself in $\mathbf{P}_{m'}$ for some $m' \leqslant m$. Thus we can use inclusion-exclusion identities to write χ_{H_m} as a linear combination of the characteristic functions of subspaces in $\mathbf{P}_{m'}$ for $m' \leqslant m$. Fortunately the resulting formula is quite simple:

Proposition 2. *The characteristic function of H_m equals*

$$\sum_{K \in \mathbf{P}_m} \chi_K \; - \; q \sum_{K \in \mathbf{P}_{m-1}} \chi_K, \tag{24}$$

in which χ_K is the characteristic function of the set K, and the second sum is interpreted as zero when $m = 0$.

Proof: Clearly (24) is an integer-valued function on W_n supported on H_m. Thus we need only show that its value at x equals 1 for all $x \in H_m$. But this value is

$$\begin{aligned} & \frac{\#(I_x \cap V_m) - 1}{q-1} \; - \; q \, \frac{\#(I_x \cap V_{m-1}) - 1}{q-1} \\ = \quad & 1 + \frac{\#(I_x \cap V_m) - q\,\#(I_x \cap V_{m-1})}{q-1} \; . \end{aligned} \tag{25}$$

Let m_0 be the degree of the minimal linear recursion satisfied by x. By Prop. 1, $I_x \cap V_m$ and $I_x \cap V_{m-1}$ are vector spaces of dimensions $m - m_0 + 1$ and $m - m_0$ respectively over k. (Note that this remains true if $m_0 = m$, when $I_x \cap V_{m-1}$ is the zero space.) Thus $\#(I_x \cap V_m) = q\;\#(I_x \cap V_{m-1})$, and (25) simplifies to 1 as claimed. ■■

We easily deduce the formula [2, Thm. 1] for the size of H_m:

Corollary. *For all nonnegative $m \leqslant (n+1)/2$ we have*

$$\#(H_m) = q^{2m}. \tag{26}$$

Proof: The size of H_m is the sum of $\chi_{H_m}(x)$ over $x \in W_n$. By (24), this sum is

$$\sum_{K \in \mathbf{P}_m} \#(K) \; - \; q \sum_{K \in \mathbf{P}_{m-1}} \#(K). \tag{27}$$

But by Lemma 2(iii), each $K \in \mathbf{P}_m$ has size q^m, and each $K \in \mathbf{P}_{m-1}$ has size q^{m-1}. Using (23) — and this is where we use the validity of (23) also for $m = -1$ — we thus simplify (27) to

$$\frac{q^{m+1}-1}{q-1}\, q^m \; - \; q\, \frac{q^m - 1}{q-1}\, q^{m-1} = q^{2m}, \tag{28}$$

as claimed. ■

In particular, if $n = 2m-1$ then $\#H_m = \#W_n$, whence $H_m = W_n$ — which is clear because in this case the Hankel matrix (2) has only m rows, so must have rank at most m. (This is essentially the special case $n = 2m-1$ of the dimension count we used earlier to deduce (13); in this case we find that $I_x \cap V_m$ has rank at least $2m - n = 1$, so must contain a nonzero vector.) Starting from this, one may establish without too much difficulty a bijection from W_{2m-1} to the subset H_m of W_n for any $n \geqslant 2m-1$, even without our $k[Y,Z]$ framework. (This is in effect how (26) is proved in [2].) But our approach also yields a formula for the Fourier transform $\widehat{\chi}_{H_m}(P)$ for all $P \in V_n$, whereas (26) only gives $\widehat{\chi}_{H_m}(0)$. We turn to $\widehat{\chi}_{H_m}$ next.

Discrete Fourier transform. To define the Fourier transform on W_n, we first define it on k. Fix a nontrivial character ψ_0 of k, that is, a nontrivial homomorphism from the additive group of k to the unit circle in $\mathbf{C}$. [If $k = \mathbf{Z}/p\mathbf{Z}$ for some prime p, we may take $\psi_0(x) = \exp(2\pi i x/p)$; in general k contains $\mathbf{Z}/p\mathbf{Z}$ where p is the characteristic of k, and we may take $\psi_0(x) = \exp(2\pi i t(x)/p)$ where $t : k \to \mathbf{Z}/p\mathbf{Z}$ is any nontrivial homomorphism of additive groups. One common choice for t is the trace from k to $\mathbf{Z}/p\mathbf{Z}$. At any rate none of our results will depend on the choice of ψ_0.] For any function $f : k \to \mathbf{C}$, we define the (discrete) *Fourier transform* $\widehat{f}$ of f to be the following function from k to $\mathbf{C}$:

$$\widehat{f}(a) := \sum_{x \in k} f(x)\psi_0(ax). \tag{29}$$

It is known that $f \mapsto \widehat{f}$ is a linear bijection on the space $\mathbf{C}^q$ of complex-valued functions on k, and that the inverse bijection is given by the *Fourier inversion formula*:

$$f(x) = \frac{1}{q} \sum_{a \in k} \widehat{f}(a)\psi_0(-ax). \tag{30}$$

The Fourier transform is defined more generally for finite-dimensional vector spaces over k. Let V, W be a dual pair of such spaces, of dimension d. (We shall use $V = V_n$, $W = W_n$, $d = n+1$.) To each function $F : W \to \mathbf{C}$ we associate its discrete Fourier transform

$$\widehat{F}(a) := \sum_{x \in W} F(x)\psi_0(\langle a, x\rangle). \tag{31}$$

Again $F \mapsto \widehat{F}$ is a linear bijection, and in this context the inversion formula reads

$$F(x) = \frac{1}{q^d} \sum_{a \in V} \widehat{F}(a)\psi_0(-\langle a, x\rangle). \tag{32}$$

To recover $\widehat{\chi}_{H_m}$ from Prop. 2, we shall need one more fact about the discrete Fourier transform:

Lemma 5. *For any linear subspace $K \subseteq W$, the Fourier transform of its characteristic function χ_K is $(\#K) \cdot \chi_{K^\perp}$, where $K^\perp$ is the annihilator of K in V.*

Proof: By definition, $\widehat{\chi}_K(a)$ is the sum over K of the character $x \mapsto \psi_0(\langle a, x\rangle)$; thus $\widehat{\chi}_K(y) = \#K$ or 0 according as this character is trivial or nontrivial on K, that is, according as $a \in K^\perp$ or $a \notin K^\perp$. ■

We can now give our formula for $\widehat{\chi}_{H_m}$. It will be convenient to introduce the following notation: for $P \in V_n$ and any integer d, define $\omega_d(P)$ to be $1/(q-1)$ times the number of nonzero $Q \in V_d$ such that P is a multiple of Q. Equivalently, $\omega_d(P)$ is the number of degree-d factors of P up to k^* scaling, and the number of homogeneous principal ideals in $k[Y, Z]$ that contain P and have a generator of degree d. For instance, $\omega_0(P) = 1$, and for all $d \geq -1$,

$$\omega_d(0) = \frac{q^{d+1}-1}{q-1} \; [= \#(\mathbf{P}_d) \text{ if } 2d \leqslant n+1]. \tag{33}$$

Moreover, for nonzero P we have the identity

$$\omega_d(P) = \omega_{n-d}(P), \tag{34}$$

due to the bijection $Q \leftrightarrow P/Q$ between factors of P of degree d and $n-d$. (The notation ω_d is suggested by the omega function in elementary number theory, which counts the positive divisors of a given positive integer.)

Theorem 1. *For every $m \leqslant (n+1)/2$ and $P \in V_n$ we have*

$$\widehat{\chi}_{H_m}(P) = q^m \left(\omega_m(P) - \omega_{m-1}(P)\right). \tag{35}$$

Proof: By Prop. 2 and Lemma 5, this follows from the following observation: for any homogeneous polynomial Q of degree at most n, the annihilator in V_n of $\ker M_n^*(Q)$ is the image of $M_n(Q)$, which is the space of degree-n multiples of Q. Thus when we use (24) to expand $\widehat{\chi}_{H_m}$ as a linear combination of characteristic functions of annihilators, the number of subspaces in $\mathbf{P}_m$ or $\mathbf{P}_{m-1}$ that contribute a term to $\widehat{\chi}_{H_m}(P)$ is the number of divisors of P of degree m or $m-1$ up to k^* scaling. Each of these terms is q^m or $-q \cdot q^{m-1} = -q^m$ respectively, whence the formula (35). ■■

As promised, Prop. 2 is the special case $P = 0$ of this formula (cf. (33)). Also, if $n = 2m-1$, the identity (34) yields $\widehat{\chi}_{H_m}(P) = 0$ for all $P \neq 0$, consistent with $H_m = W_n$ in that case.

Hankel matrices with independently biased entries. The formula (26) can be interpreted thus: if $x_0, \ldots, x_n$ are chosen independently at random from the uniform distribution on k, then the resulting vector $(x_0, \ldots, x_n)$ is in H_m with probability $q^{2m-(n+1)}$. Using Thm. 1 we can also get at the probability that $(x_0, \ldots, x_n) \in H_m$ if the x_i are still chosen independently at random but from distributions μ_i on k that are not necessarily uniform.

We regard the μ_i as functions from k to $\mathbf{R}$ satisfying the conditions: $\mu_i(x) \geqslant 0$ for all $x \in k$, and

$$[\widehat{\mu}_i(0) =] \sum_{x \in k} \mu_i(x) = 1. \tag{36}$$

Then the probability that $\vec{x} := (x_0, \ldots, x_n)$ is in H_m is

$$\Pi_m(\mu_0, \ldots, \mu_n) = \sum_{\vec{x} \in W_n} \chi_{H_m}(\vec{x}) \prod_{i=0}^{n} \mu_i(x_i). \tag{37}$$

By applying Fourier inversion to χ_{H_m} we can express this as a linear combination of the values of $\widehat{\chi}_{H_m}(P)$. The resulting formula is:

Lemma 6. *We have*

$$\Pi_m(\mu_0, \ldots, \mu_n) = q^{-(n+1)} \sum_{P \in V_n} \widehat{\chi}_{H_m}(P) \prod_{i=0}^{n} \widehat{\mu}_i(-a_i), \tag{38}$$

where a_i is the $Y^i Z^{n-i}$ coefficient of P as in (4).

Proof: By Fourier inversion (32),

$$\Pi_m(\mu_0, \ldots, \mu_n) = q^{-(n+1)} \sum_{P \in V_n} \widehat{\chi}_{H_m}(P) \left(\sum_{\vec{x} \in W_n} \psi_0(-\langle P, x \rangle) \prod_{i=0}^{n} \mu_i(x_i) \right). \tag{39}$$

Now $\langle P, x \rangle = \sum_{i=0}^{n} a_i x_i$, so

$$\psi_0(-\langle P, x \rangle) \prod_{i=0}^{n} \mu_i(x_i) = \prod_{i=0}^{n} \psi_0(-a_i x_i) \mu_i(x_i). \tag{40}$$

Thus the inner sum in (39) factors into

$$\prod_{i=0}^{n} \left(\sum_{x_i \in k} \psi_0(-a_i x_i) \mu_i(x_i) \right) = \prod_{i=0}^{n} \widehat{\mu}_i(-a_i). \tag{41}$$

Entering this into (39) yields the claimed formula (38). ■

The term $P = 0$ in (39) contributes

$$q^{-(n+1)}\widehat{\chi}_{H_m}(0)\prod_{i=0}^{n}\widehat{\mu}_i(0) = q^{2m-(n+1)}, \tag{42}$$

because $\widehat{\chi}_{H_m}(0) = q^{2m}$ and each $\widehat{\mu}_i(0) = 1$. The absolute value of the sum of the remaining terms is at most

$$\begin{aligned} & q^{-n+1} \sup_{P\in V_n-\{0\}} |\widehat{\chi}_{H_m}(P)| \cdot \sum_{P\in V_n-\{0\}} \prod_{i=0}^{n} |\widehat{\mu}_i(-a_i)| \\ = \; & q^{-n+1} \sup_{P\in V_n-\{0\}} |\widehat{\chi}_{H_m}(P)| \left[\left(\prod_{i=0}^{n} \|\widehat{\mu}_i\|_1\right) - 1\right], \end{aligned} \tag{43}$$

where $\|\widehat{\mu}_i\|_1$ is the l_1 norm

$$\|\widehat{\mu}_i\|_1 := \sum_{a\in k} |\widehat{\mu}_i(a)|. \tag{44}$$

Since $\widehat{\mu}_i(0) = 1$, we have $\|\widehat{\mu}_i\|_1 \geqslant 1$, with equality if and only if $\widehat{\mu}_i(a) = 0$ for all $a \neq 0$. By Fourier inversion (30), this condition is equivalent to $\mu_i(x) = 1/q$ for all x. Hence $\|\widehat{\mu}_i\|_1 = 1$ if and ony if μ_i is the uniform distribution on k. We may thus regard $(\prod_{i=0}^{n} \|\widehat{\mu}_i\|_1) - 1$ as a measure of how far the product distribution $\mu_0 \cdots \mu_n$ departs from uniform distribution on W_n.

What of the other factor $\sup_{P\neq 0} |\widehat{\chi}_{H_m}(P)|$ in the error estimate (43)? By Thm. 1, each $\widehat{\chi}_{H_m}(P)$ is a multiple of q^m. Once $n \geqslant 2m$, we cannot expect $\widehat{\chi}_{H_m}(P)$ to vanish for all $P \neq 0$, so $\sup_{P\neq 0} |\widehat{\chi}_{H_m}(P)|$ must be at least q^m. We next show that it $|\widehat{\chi}_{H_m}(P)|$ is never much larger than q^m for $P \neq 0$:

Lemma 7. *For every q and $\epsilon > 0$, there exists an effective constant C such that*

$$\omega_d(P) < C(1+\epsilon)^n \tag{45}$$

for every nonzero $P \in V_n$ and every integer d.

(This is analogous to the standard fact that the number of factors of an n-digit integer is subexponential in n, and will be proved in the same way.)

Proof: Define

$$\omega(P) := \sum_{d=0}^{n} \omega_d(P), \tag{46}$$

the total number of divisors of P up to k^* scaling. Factor P into irreducibles over k:

$$P = \prod_{s=1}^{r} P_s^{e_s}, \tag{47}$$

with P_s distinct irreducibles of degree f_s. Comparing degrees in (47) we find

$$n = \sum_{s=1}^{r} e_s f_s. \tag{48}$$

Now

$$\omega(P) = \prod_{s=1}^{r} (e_s + 1), \tag{49}$$

because the general divisor of P is $\prod_{s=1}^{r} P_s^{e'_s}$ with each e'_s chosen from among the $e_s + 1$ possibilities $0, 1, \ldots, e_s$. Fix m_0 large enough that $2^{1/m_0} < 1 + \epsilon$, and factor (49) as

$$\omega(P) = \prod_{f_s < m_0} (e_s + 1) \prod_{f_s \geqslant m_0} (e_s + 1). \tag{50}$$

The second product is at most

$$\prod_{f_s \geqslant m_0} 2^{e_s} = 2^{\sum_{f_s \geqslant m_0} e_s} \leqslant 2^{n/m_0}, \tag{51}$$

since $m_0 \sum_{f_s \geqslant m_0} e_s \leqslant \sum_{s=1}^{r} e_s f_s = n$ by (48). The first product in (50) has at most B factors, where B is the number of irreducible bivariate homogeneous polynomials of degree $< m_0$ up to k^* scaling. Each factor is at most $n + 1$, so the product is at most $(n + 1)^B$. Since $\log(n + 1)^B = o(n)$ as $n \to \infty$, and $2^{1/m_0} < 1 + \epsilon$, we conclude that

$$\omega(P) \leqslant 2^{n/m_0} (n + 1)^B \ll (1 + \epsilon)^n. \tag{52}$$

Since $\omega_d(P) \leqslant \omega(P)$, we deduce $\omega_d(P) \ll (1 + \epsilon)^n$. ■

Combining this estimate with Thm. 1 and Lemma 6, we obtain:

Theorem 2. *For every q and $\epsilon > 0$, there exists an effective constant C such that*

$$\left| \Pi_m(\mu_0, \ldots, \mu_n) - q^{2m-(n+1)} \right| < C(1 + \epsilon)^n q^{m-n} \prod_{i=0}^{n} \|\widehat{\mu}_i\|_1 . \tag{53}$$

for any n and any distributions μ_i on k. ■■

In particular, suppose that $n = 2m + \alpha$ for some fixed nonnegative integer α, and that all the μ_i are the same, so that each x_i is chosen from the same distribution μ. Then, as long as $\|\widehat{\mu}\|_1 < q^{1/2}$, the error term in (53) approaches 0 as $m \to \infty$, and we conclude that if each of $x_0, \ldots, x_{2m+\alpha}$ is chosen independently from the distribution μ then $x \in H_m$ with probability approaching $q^{-(\alpha+1)}$, same as for the uniform distribution. As noted in the Introduction, the bound on $\|\widehat{\mu}\|_1$ is best possible, at least if q is a square: in that case k has a quadratic subfield k_0, and if each x_i is chosen uniformly from k_0 (or from ck_0 for some $c \in k^*$) then $x \in H_m$ with probability $q^{-(\alpha+1)/2}$, not $q^{-(\alpha+1)}$; but for this distribution, $\|\widehat{\mu}\|_1 = q^{1/2}$ by Lemma 5.

Open questions

Better bounds on $\Pi_m(\mu_0,\ldots,\mu_n) - q^{2m-(n+1)}$? We showed (Thm. 2) that $\Pi_m(\mu_0,\ldots,\mu_n)$ is well approximated by $q^{2m-(n+1)}$ under certain hypotheses on the μ_i. Can these hypotheses by weakened by lowering the error bound in (53)? Of course we must exclude some choices of μ_i. For instance we certainly cannot have every μ_i supported on only one point; and we already gave the counterexample of uniform distribution on a proper subfield of k. But it seems plausible that, except for such pathological cases, $(x_0,\ldots,x_n)$ should be about as likely to be in H_m with x_i chosen from μ_i as it is with $\vec{x}$ chosen uniformly from W_n — whether or not the $\|\mu_i\|_1$ are small enough to deduce $\Pi_m(\mu_0,\ldots,\mu_n) \sim q^{2m-(n+1)}$ from Thm. 2. For instance we may surmise the following

Conjecture. *Fix k and a closed set K of distributions $\mu : k \to \mathbf{R}$. Assume that no $\mu \in K$ is supported on a single point, nor on ck_0 for any $c \in k^*$ and any proper subfield k_0 of k. Then, for every real $R \geqslant 2$, we have*

$$\Pi_m(\mu_0,\ldots,\mu_n) = (1+o(1))q^{2m-(n+1)} \tag{54}$$

for any sequence of $(n, m, \mu_0,\ldots,\mu_n)$ for which $m \to \infty$, $2m \leqslant n \leqslant Rm$, and $\mu_i \in K$ for each i.

In particular, suppose $q = R = 2$. A distribution on k is then a pair $(\mu(0), \mu(1))$ of nonnegative numbers with $\mu(0)+\mu(1) = 1$. The conjecture then asserts that, for each $p > 0$, if each entry x_i of a square Hankel matrix of order $m+1$ over $\mathbf{Z}/2\mathbf{Z}$ is chosen independently at random with probabilities $\mu_i(0), \mu_i(1)$ both $\geqslant p$, then the matrix is singular with probability approaching $1/2$ as $m \to \infty$. Thm. 2 shows this only for $p > 1 - 2^{-1/2} \approx 29.3\%$.

Higher dimensions. What happens to our theory in the context of arrays of dimension 2 or greater, rather than finite sequences? One could start the analysis in the same way, using for instance homogeneous polynomials in three variables to treat triangular arrays, or bihomogeneous polynomials in two pairs of variables for rectangular arrays. The resulting structures will surely be more complicated in higher dimensions, but it may still be possible to find tractable descriptions.

Determinants of nonsingular Hankel matrices. In another direction, we return to the case $n = 2m$ of square Hankel matrices (3) of order $m+1$, for which H_m consists in effect of such matrices whose determinant vanishes. We then ask: is there a formula analogous to (35), or even an estimate analogous to Lemma 7, for the discrete Fourier transform of the set of square Hankel matrices of order $m+1$ with determinant c, for any given *nonzero* $c \in k$? This is easy when $q = 2$, in which case that set is just the complement of H_m. But the problem seems to require new techniques once $q \geqslant 3$.

References

[1] Cantor, D.G.: On the analogue of the division polynomials for hyperelliptic curves. *J. Reine Angew. Math.* (*Crelle's J.*) **447** (1994), 91–145.

[2] Daykin, D.: Distribution of bordered persymmetric matrices in a finite field. *J. Reine Angew. Math.* (*Crelle's J.*) **203** (1960), 47–54.

[3] Dress, A., Elkies, N.D., Luca, F.: A characterization of Mahler's generalized Liouville numbers by simultaneous rational approximation. Preprint, 2001.

[4] Iohvidov, I.S.: *Hankel and Toeplitz Matrices and Forms: Algebraic Theory* (trans. G.P.A. Thijsse). Boston: Birkhäuser 1982.